心理学报 2023, Vol. 55, No. 8, 1220−1233
Acta Psychologica Sinica

https://doi.org/10.3724/SP.J.1041.2023.01220

返回抑制对视听跨通道对应的影响*

祖光耀[1] 李舒淇[1] 张天阳[2] 王爱君[1] 张 明[1]

([1]苏州大学心理学系, 心理与行为科学研究中心; [2]苏州大学公共卫生学院, 江苏 苏州 215123)

摘 要 视听跨通道对应(audiovisual crossmodal correspondence)在不同类型的视、听刺激之间被广泛发现, 但其发生阶段尚不明确。研究采用线索−靶子范式, 探究返回抑制(inhibition of return, IOR)对于视听跨通道对应的影响。实验 1 操纵线索和目标之间的空间位置一致性和听觉音调与视觉目标位置之间的跨通道对应一致性, 结果发现 IOR 效应和视听跨通道对应存在交互作用, 线索化位置存在稳定的视听跨通道对应效应, 而非线索化位置的视听跨通道对应效应消失。实验 2 操纵无关听觉刺激是否出现, 发现 IOR 效应与单独的声音出现与否不存在交互作用, 排除了警觉效应对于结果的混淆。实验 3 通过延长线索和目标的间隔时间(stimulus onset asynchronies, SOA), 发现伴随着 IOR 效应减弱, 线索化位置的视听跨通道对应效应也相应减弱, 且 IOR 效应对于视听跨通道对应的调节作用减弱。研究表明, 仅当听觉刺激与视觉空间位置之间发生跨通道对应时, 才会与同样发生在知觉水平的 IOR 效应产生交互, 支持了视听跨通道对应发生在知觉阶段。同时研究结果支持视听跨通道对应的发生符合反比效应原则。

关键词 视听跨通道对应, 返回抑制, 线索−靶子范式, 警觉效应

分类号 B842

1 引言

人类依靠多个感觉通道去感受周围的坏境, 不同感觉通道间信号的结合会促进人类的行为反应, 这一现象被称为多感觉反应增强(Frassinetti et al., 2002; Stein et al., 1989)。以往对于多感觉反应增强的研究集中于多感觉整合(McCracken et al., 2019; Starke et al., 2017)。多感觉整合是指当不同感觉通道的信息具有时空邻近性时, 个体将这些信息进行整合以形成连贯和有意义的表征(彭姓 等, 2019; Tang et al., 2016)。视觉和听觉输入信息的整合被称为视听觉整合(audiovisual integration), 相比于单一视觉或听觉, 视、听双通道的整合会产生冗余效应(redundant effect), 促进被试对于双通道刺激的探测和辨别(彭姓 等, 2019; 唐晓雨 等, 2020; Stein & Stanford, 2008; Talsma & Woldorff, 2005)。如相同位置同时呈现的视觉和听觉刺激提供了相同的空间和时间信息, 提高了个体的反应能力(Spence, 2013)。

以往研究发现, 视、听刺激除了以一种冗余的方式影响被试的反应外, 还能以一种非冗余的方式对被试的反应产生影响, 即视、听刺激提供了两种不同维度的信息, 但却相互映射影响当前的行为, 这种现象被称为视听跨通道对应(audiovisual crossmodal correspondence) (Spence, 2011, 2019)。一种常见的视听跨通道对应是听觉的音调与视觉的空间位置之间的对应, 个体倾向于将音调较高的声音与高空间位置相映射, 而将音调较低的声音与低空间位置相映射。当高音伴随或者先于视觉刺激呈现时, 被试对于高空间位置呈现的视觉刺激的反应快于低空间位置呈现的视觉刺激, 反之亦然(Chiou & Rich, 2012; Evans & Treisman, 2010; McCormick et

---

收稿日期: 2022-9-27

* 教育部人文社会科学研究项目(22YJCZH243)、苏州市医疗卫生科技创新项目(SKY2022113)、国家自然科学基金项目(31800907)、江苏高校哲学社会科学研究项目(2022SJYB1454)和苏州大学人文社科青年交叉团队(2022)资助。

李舒淇和祖光耀同为第一作者

通信作者: 王爱君, E-mail: ajwang@suda.edu.cn; 张天阳, E-mail: tyzhang@suda.edu.cn

al., 2018; Spence, 2019; Zeljko et al., 2019)。除此之外，音调与刺激大小(Brunetti et al., 2018; Parise & Spence, 2008)、音调与亮度(Maimon et al., 2020; Marks et al., 2003)以及音调与空间频率(Evans & Treisman, 2010)都存在视听跨通道对应。不同于视听觉整合所需要的视觉和听觉刺激在时间和空间上相邻呈现(Spence, 2013)，视听跨通道对应发生时，视觉刺激和听觉刺激的呈现位置可以不同，呈现的时间间隔也可以相对更长(Chiou & Rich, 2012)。视听跨通道对应是一种相对的映射(Chiou & Rich, 2012)，如在音调与空间位置的跨通道对应中，不存在某个频率的音调与某个高度的空间位置之间的绝对的映射，而是存在两个音调中相对较高或较低的音调与两个位置中相对较高或较低的位置之间的映射。

现有研究针对视听跨通道对应的发生阶段存在不同观点。第一种观点认为，视听跨通道对应发生在知觉水平，即视听跨通道对应增加了刺激的感知觉显著性(Evans & Treisman, 2010)。已有研究发现当被试观看模糊移动的光栅时，如果伴随着音调的上升，被试更倾向于知觉到光栅在向上运动，如果伴随着音调的下降，被试更倾向于知觉到光栅在向下运动(Maeda et al., 2004)。ERP (event-related potential)研究发现当先前呈现的声音与当前视觉符号存在对应一致时，视觉诱发的早期知觉成分N1 的波幅相比于对应不一致时更大，支持了视听跨通道对应是知觉上的增强(Ković et al., 2010)。第二种观点认为，视听跨通道对应发生在语义层面(Spence, 2011)。研究者发现在几乎所有的语言环境中，人们都用“高” “低”来描述不同频率的音调，而同样的词语又对应空间位置的高和低，以及空间频率的高和低，因此研究者认为视、听不同维度刺激相同的语义编码导致了视听跨通道对应的出现(Walker, 2012)。研究者用“高” “低”两个词的语音代替高、低频率的音调，发现了语音与空间位置也出现了跨通道对应，支持了视听跨通道对应发生在语义层面的观点(Gallace & Spence, 2006)。第三种观点认为，视听跨通道对应发生在晚期决策水平，即一个通道与另一个通道的对应降低了对于所呈现目标刺激的反应标准(Spence, 2011)。研究者使用信号检测论的方法对于视听跨通道对应进行研究，结果显示当视听跨通道对应发生时，被试的知觉敏感性 *d’* 没有发生变化，而对于目标的虚报率则上升。研究者认为如果视听跨通道对应发生在知觉层面，被试反应能力的增强不会伴随虚报率的上升，因而视听跨通道对应发生在决策而非知觉水平(Marks et al., 2003)。总之，由于所选用的刺激材料和测量指标不同，视听跨通道对应的发生阶段尚无定论(Spence, 2011)。

视、听刺激的整合或对应可以提高目标的知觉显著性(Evans & Treisman, 2010; Ković et al., 2010; Tang et al., 2019)，促进个体的行为反应，而注意系统中的返回抑制(inhibition of return, IOR)也会影响人类对于目标的知觉(唐晓雨 等, 2020; Tang et al., 2019)。IOR 是指在线索−靶子范式(cue-target paradigm)中，当线索和目标之间的 SOA 大于约 300 ms 时，被试对于线索化位置出现的刺激的反应时慢于非线索位置出现的刺激的现象(Posner & Cohen, 1984)。IOR 避免了相同位置被重复搜索，提高了视觉搜索效率(Redden et al., 2021)。尽管对于 IOR 的发生机制存有不同的理论解释，但人们普遍认为IOR 反映了一种知觉抑制，即注意从线索化位置脱离，导致线索化位置上目标的感知觉显著性降低，使得反应受到损害(Klein, 2000; Satel et al., 2013)。ERP 研究发现 IOR 发生时线索化位置刺激诱发的早期 P1 和 N1 成分波幅低于非线索化位置(Hopfinger & Mangun, 2001; Prime & Jolicoeur, 2009)，支持了IOR 效应发生在早期的感知觉阶段。以往研究关注IOR 效应和多感觉刺激间的相互作用，这样的研究集中于视听觉整合(彭姓 等, 2019; Tang et al., 2019; van der Stoep, van der Stigchel, et al., 2015)。研究者在线索−靶子范式中设置视听觉目标，结果发现IOR 效应对于视听觉整合产生了调节作用。一部分研究者发现线索化位置的视听觉整合效应更小(彭姓 等, 2019; Tang et al., 2019; van der Stoep et al., 2016)，另一部分研究者发现了相反的结果，即线索化位置的视听觉整合效应更大(唐晓雨 等, 2020)。不同的研究结果可能与不同实验的 SOA 设置有关(唐晓雨 等, 2020)，但已有研究都表明，视听觉整合发生在知觉阶段(Tang et al., 2019)，因此会受到同样发生在知觉加工阶段的 IOR 效应的调节(彭姓 等, 2019; 唐晓雨 等, 2020)。

尽管视听觉整合和视听跨通道对应以不同的方式导致多感觉反应增强，前者通过时间和空间上的临近提供冗余的信息(Noesselt et al., 2007; Santangelo et al., 2008)，后者通过不同维度信息间的相互映射促进当前的行为(Chiou & Rich, 2012; McCormick et al., 2018)。但如果视听跨通道对应发

生在知觉阶段，结合以往研究发现的多感觉刺激与包括 IOR 在内的注意线索效应之间的相互作用(唐晓雨 等, 2020; Botta et al., 2017; Chica et al., 2011; Tang et al., 2019)，以及加因素法反应时的实验逻辑(Sternberg, 1969)，则 IOR 效应发生时，线索化位置目标的感知觉显著性降低会对于视听跨通道对应产生影响；而如果视听跨通道对应发生在语义或者决策水平，则 IOR 效应不会对于视听跨通道对应产生影响。因此，本研究将空间线索范式与视听跨通道对应的研究范式相结合，对于 IOR 效应和视听跨通道对应之间的关系进行了探究。以往视听跨通道对应的研究中，视、听刺激同步呈现(Brunel et al., 2015; Gallace & Spence, 2006; Getz & Kubovy, 2018)，导致所测量到的结果可能同时包含了视听觉整合和视听跨通道对应。考虑到视、听刺激呈现的时空邻近性是视听觉整合发生的必要条件(Spence, 2011)，已有研究表明当视、听刺激呈现间隔时间超过 100 ms 时视听觉整合效应消失(van der Stoep, Spence, et al., 2015)，因此本研究在听觉刺激消失后 200 ms 呈现视觉目标，且听觉刺激以双声道形式呈现，以最小化视听觉整合的影响。

另外，先于视觉目标呈现的听觉刺激会对于目标产生警觉效应(Wiegand & Sander, 2019)，这种警觉效应会提高被试对于视觉刺激的知觉能力(Kusnir et al., 2011)，并与目标的感知觉显著性产生交互作用(Botta et al., 2017)，因此目前研究中警觉效应可能与外源性线索引发的 IOR 效应相互作用，造成不同条件下的警觉效应不同。考虑到目前研究是第一个探究 IOR 效应与视听跨通道对应之间关系的研究，在实验中排除可能存在的混淆因素是必要的，因此本研究设计了控制实验，排除目前范式中警觉效应与 IOR 效应产生交互进而混淆实验结果的可能。最后，为了进一步探究 IOR 效应和视听跨通道对应之间交互作用的机制，本研究通过操纵线索和目标之间的 SOA 来操纵 IOR 效应的大小(Lupiáñez et al., 1997)，进而探究 IOR 效应对于视听跨通道对应的调节机制。

综上，本研究采用线索−靶子范式，在视觉目标之前呈现听觉刺激，考察 IOR 对于视听跨通道对应的影响。研究分为 3 个实验，实验 1 操纵线索和目标之间的空间线索有效性和听觉刺激与视觉目标之间的跨通道对应一致性，探究 IOR 效应与视听跨通道对应之间的关系。实验 1 假设，音调与空间位置之间的视听跨通道对应发生在知觉水平，因此会受到发生在相同加工阶段的 IOR 效应的影响，出现视听跨通道对应与 IOR 效应之间的交互作用。实验 2 通过操纵听觉刺激是否出现，探究 IOR 效应与单独的声音出现与否之间的关系。实验 2 中听觉刺激只存在一种音调，而音调与空间位置之间的跨通道对应需要呈现两个具有高低关系的音调才能产生(Chiou & Rich, 2012)，因此只呈现一种音调不会与视觉目标的空间位置产生跨通道对应。实验 2 的目的一方面是验证在目前的研究中，在视觉刺激前单纯呈现听觉刺激不会与 IOR 效应产生交互，而只有听觉刺激与视觉目标发生跨通道对应才会与 IOR 效应产生交互；另一方面是排除可能存在的警觉效应对于结果的混淆。基于以往研究表明，警觉效应是以一种自上而下的方式增强对于刺激的感知(Kusnir et al., 2011)，而 IOR 以自下而上的方式影响被试的知觉(Berdica et al., 2017; Jia et al., 2019)，实验 2 假设 IOR 效应不会与听觉刺激的出现与否产生交互，进一步支持实验 1 的结果是由 IOR 效应对视听跨通道对应的影响所导致的。实验 3 通过操纵线索和目标之间的 SOA 来操纵 IOR 效应的大小，探究 IOR 效应对于视听跨通道对应的调节机制。根据多感觉反应增强中的反比效应原则(principle of inverse effectiveness) (Meredith & Stein, 1983; van der Stoep et al., 2016)，研究预期随着线索和目标之间的 SOA 增加，IOR 效应变小(Lupiáñez et al., 1997)，进而使得线索化位置的视听跨通道对应效应变小，且 IOR 效应对于视听跨通道对应的调节作用减弱。

2 实验 1：IOR 效应对视听跨通道对应的影响

2.1 方法

2.1.1 被试

采用 G*Power 3.1 软件对样本量进行计算(Erdfelder et al., 2009; Faul et al., 2007)。设置 I 类错误的概率 α error prob 为 0.05，检验效能 Power (1 − β err prob)为 0.8，效应量为中等水平(f = 0.25) (Cohen, 1992)，计算被试样本量为 24。招募江苏省高校在校学生 31 名，其中男生 14 名，女生 17 名，年龄 18~24 岁。所有被试均为右利手，听力、视力或矫正视力正常，无神经或精神疾病史，无脑部损伤史。被试在实验完成后会获得相应的报酬。

2.1.2 实验仪器与材料

实验程序使用 E-prime 2.0 编制，运行于 Dell

3020 MT 主机上。实验刺激呈现在 23 英寸的液晶显示器(Dell E2316Hf)上, 屏幕的分辨率为 1024 × 768, 刷新率为 60 Hz。整个实验中被试头部通过下巴托固定, 距离屏幕 60 cm。实验在隔音的暗环境下进行。

所有的视觉刺激材料都在白色的背景上由黑色(RGB: 0, 0, 0)线条绘制。在每个试次中, 3 个正方形方框(1.5°× 1.5°)竖直排列在屏幕上, 其中一个方框位于屏幕正中, 另外两个方框置于上下两侧, 两个相邻方框相距视角为 4.5°。中央注视点(1° × 1°)呈现在中央方框内, 线索通过对于注视点上方或者下方的矩形边框加粗 0.5°实现, 中央线索通过中央注视点变大为 1.5° × 1.5°实现。视觉目标为一个圆盘(1° × 1°), 听觉刺激为 250 Hz 或 2500 Hz 的正弦音(50 ms)。声音刺激通过头戴式铁三角耳麦(ATH-WS99)从左右耳双声道呈现, 音量控制在 65 dB。

2.1.3　实验设计与流程

实验 1 采用 2 (空间线索有效性: 线索有效 vs. 线索无效) × 2 (跨通道对应一致性: 对应一致 vs.对应不一致)的被试内设计, 因变量为反应时和正确率。其中, 跨通道对应一致条件是指在高音后呈现高位置的视觉目标, 在低音后呈现低位置的视觉目标; 跨通道对应不一致条件则与之相反。正式实验一共 5 组, 每组 53 个试次, 其中包含 5 个捕获试次, 共 265 个试次。被试在进行正式实验前完成 53 个试次的练习实验, 整个实验大约进行 35 分钟。

实验 1 中每个试次的流程如图 1 所示。首先, 屏幕中间会呈现注视点“+”, 呈现时间为 750 ms。随后注视点上方或者下方的方框边框会加粗, 作为线索, 呈现时间为 50 ms, 线索对于目标的空间位置没有预测性。在 250 ms 的时间间隔后, 注视点加粗作为中央线索呈现。中央线索是空间 IOR 研究中常用的设置, 有助于 IOR 效应更稳定的出现(Prime et al., 2006)。中央线索呈现 50 ms, 之后随机呈现不同音调的听觉刺激。听觉刺激呈现时间为 50 ms, 在 200 ms 的时间间隔后, 视觉目标在注视点上方或者下方的方框内呈现 100 ms。被试需要在探测到视觉目标时进行既快又准的按键反应, 捕获试次不需要按键。如果 1000 ms 内被试没有按键, 则自动进入下一个试次。除了练习实验, 实验中没有给予被试按键正确或错误的反馈。

2.2　结果与分析

错误反应、没有反应和反应时异常(100 ms 以内或超出平均反应时正负 3 个标准差)的数据不计入反应时分析, 剔除的数据量占总数据量的 1.09%。实验 1 为简单的探测反应任务, 被试的平均正确率达到 98%以上, 因此不针对正确率进行进一步的统计分析。

对反应时进行 2 (线索有效性: 线索有效 vs. 线索无效) × 2 (跨通道对应一致性: 对应一致 vs. 对应不一致)的重复测量方差分析。结果如图 2 所示, 线索有效性主效应显著, $F(1, 30) = 122.26$, $p < 0.001$, $\eta_p^2 = 0.80$, 线索有效条件下的反应时(325 ms)显著慢于线索无效条件(288 ms), IOR 效应出现。跨通道对应一致性主效应显著, $F(1, 30) = 4.95$, $p = 0.034$, $\eta_p^2 = 0.14$, 对应一致条件下的反应时(305 ms)显著快于对应不一致条件(308 ms), 视听跨通道对应效应出现。重要的是, 线索有效性和跨通道对应一致性的交互作用显著, $F(1, 30) = 6.69$, $p = 0.015$, $\eta_p^2 = 0.18$, IOR 效应对于视听跨通道对应产生了调节作用。简单效应分析表明, 当线索有效时, 跨通道对应一致条件下的反应时(322 ms)显著快于对应

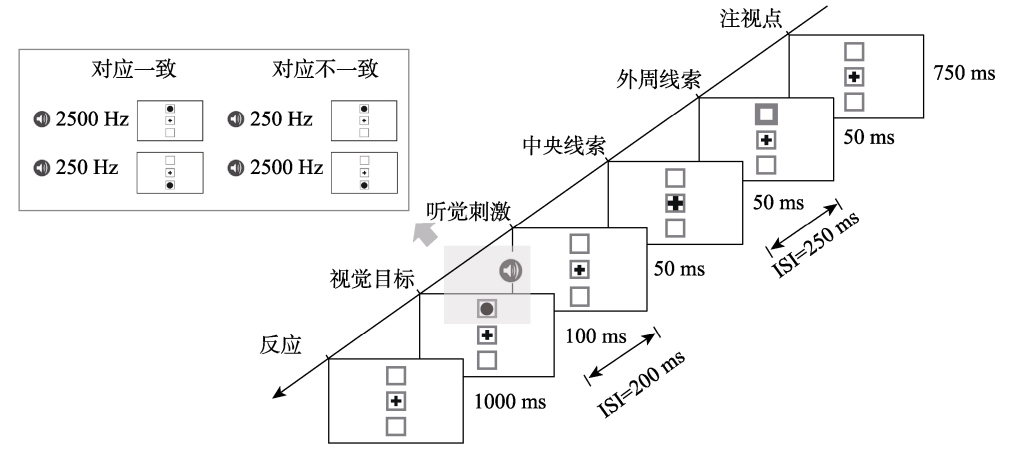


图 1　实验 1 流程图

不一致条件(327 ms), $t(30) = 3.26$, $p = 0.003$, Cohen's $d = 0.59$, 95% CI = [−9.29, −2.13], 出现了视听跨通道对应效应; 当线索无效时, 跨通道对应一致条件下的反应时(289 ms)和对应不一致条件下的反应时(288 ms)没有显著差异, $t(30) < 1$, 没有出现视听跨通道对应效应。另一个维度的简单效应分析表明, 当跨通道对应一致时, 线索有效条件下的反应时(322 ms)显著慢于线索无效条件(288 ms), $t(30) = 10.19$, $p < 0.001$, Cohen's $d = 1.83$, 95% CI = [26.76, 40.19], IOR 效应出现; 当跨通道对应不一致时, 线索有效条件下的反应时(327 ms)显著慢于线索无效条件(288 ms), $t(30) = 10.76$, $p < 0.001$, Cohen's $d = 1.93$, 95% CI = [31.79, 40.69], IOR 效应同样出现。对于 IOR 效应量(线索有效条件减去线索无效条件的平均反应时)进行配对样本 t 检验, 结果发现跨通道对应一致条件下的IOR效应量(33 ms)显著小于不一致条件(39 ms), $t(30) = 2.59$, $p = 0.015$, Cohen's $d = 0.47$, 95% CI = [−10.31, −1.21], 视听跨通道对应产生的促进作用抵消了部分 IOR 效应。

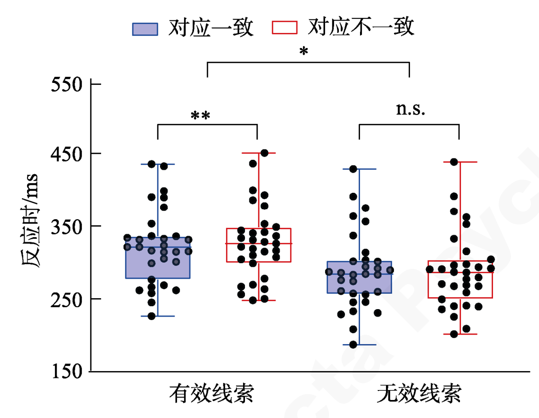


图 2 实验 1 各个条件下的箱线图

注：线条自上而下为上限、上四分位数、中位数、下四分位数、下限。散点为个体平均反应时。*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, n.s.代表无显著差异

2.3 小结

实验 1 的结果发现, 被试在线索有效条件下的反应时慢于线索无效条件, IOR 效应出现, 表明实验在中央线索后呈现听觉刺激并没有影响 IOR 效应的产生。此外, 被试在跨通道对应一致条件下的反应时快于对应不一致条件, 听觉音调与视觉空间位置之间的跨通道对应出现。重要的是, IOR 效应与视听跨通道对应产生了交互作用, 线索有效条件下出现了视听跨通道对应效应, 而线索无效条件下没有出现视听跨通道对应效应, 发生在早期知觉阶段的 IOR 效应对于视听跨通道对应产生了调节作用。

以往视听觉整合的研究中, 有研究者发现 IOR 效应发生时, 线索有效条件下的视听觉整合效应更大(唐晓雨 等, 2020)。研究者通过反比效应原则对此加以解释, 即更弱的视、听感觉信号输入引发了更强的整合(Meredith & Stein, 1983)。目前研究中 IOR 效应发生时线索化位置的目标感知觉显著性降低, 非线索化位置的目标感知显著性相对提高(Satel et al., 2013), 可能导致了视觉目标与听觉刺激之间的跨通道对应仅在线索化位置出现。这意味着反比效应原则可能也适用于视听跨通道对应, 具体将在总讨论中进一步论述。实验 1 还发现跨通道对应一致条件下的 IOR 效应量更小, 意味着视听跨通道对应部分抵消了 IOR 效应对于目标在早期感知觉上的抑制, 这与视听觉整合研究中的发现相一致(Tang et al., 2019)。

实验 1 的结果支持了听觉音调与视觉空间位置之间的跨通道对应与 IOR 效应均发生在早期的知觉阶段, 因而两者之间会产生交互。但在目前研究中, 先于视觉目标呈现的听觉刺激会产生警觉效应(Wiegand & Sander, 2019), 这种警觉效应可能在线索化位置和非线索化位置的大小不同(Botta et al., 2017), 进而对于实验结果产生混淆。为了进一步支持实验 1 的结果是由 IOR 效应对于视听跨通道对应进行调节所导致的, 本研究设计了实验 2, 实验 2 通过操纵听觉刺激是否出现, 探究 IOR 效应与单独的声音出现与否之间的关系。实验 2 的目的一方面是验证在目前的研究中, 在视觉刺激前单纯呈现听觉刺激不会与 IOR 效应产生交互, 而只有听觉刺激与视觉目标发生跨通道对应才会与 IOR 效应产生交互; 另一方面是排除可能存在的警觉效应对于结果的混淆。

3 实验 2:IOR 效应与声音呈现与否的关系

3.1 方法

3.1.1 被试

采用 G*Power 3.1 软件对样本量进行计算。设置 I 类错误的概率 α error prob 为 0.05, 检验效能 Power (1 − β err prob)为 0.8, 效应量为中等水平(f = 0.25) (Cohen, 1992), 计算被试样本量为 24。招募江苏省高校在校学生 34 名, 其中男生 15 名, 女生 19 名, 年龄 18~24 岁。所有被试均为右利手, 听力、视力或矫正视力正常, 无神经或精神疾病史, 无脑部损伤史。被试在实验完成后会获得相应的报酬。

3.1.2 实验仪器与材料

实验 2 中的听觉刺激为 1600 Hz 的正弦音，其余的实验仪器与实验材料与实验 1 相同。

3.1.3 实验设计和实验流程

实验 2 采用 2 (空间线索有效性：线索有效 vs. 线索无效) × 2 (声音呈现：有声音 vs.无声音)的被试内设计，因变量为反应时和正确率。实验 2 与实验 1 的不同之处是听觉刺激为 1600 Hz 的纯音，可能出现在视觉刺激之前，也可能不出现。其余实验流程和试次设置均与实验 1 保持一致。

3.2 结果与分析

错误反应、没有反应和反应时异常(100 ms 以内或超出平均反应时正负 3 个标准差)的数据不计入反应时分析，剔除的数据量占总数据量的 1.88%。在实验 2 中，被试的平均正确率达到 98% 以上，因此不针对正确率进行进一步的统计分析。

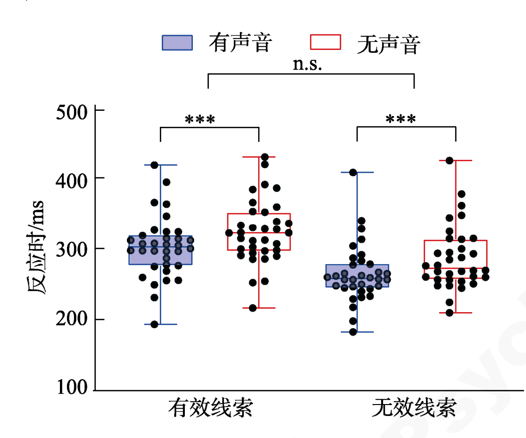


图 3　实验 2 各个条件下的箱线图

注：线条自上而下为上限、上四分位数、中位数、下四分位数、下限。散点为个体平均反应时。*p < 0.05, **p < 0.01, ***p < 0.001, n.s.代表无显著差异

对反应时进行 2 (线索有效性：线索有效 vs. 线索无效) × 2 (声音呈现：有声音 vs. 无声音)的重复测量方差分析。结果如图 3 所示，线索有效性主效应显著，$F(1, 33) = 237.78$, $p < 0.001$, $\eta_p^2 = 0.88$, 线索有效条件下的反应时(313 ms)显著慢于线索无效条件(294 ms), IOR 效应出现。声音呈现主效应显著，$F(1, 33) = 82.34$, $p < 0.001$, $\eta_p^2 = 0.71$, 有声音条件下的反应时(283 ms)显著快于无声音条件(305 ms)，听觉刺激促进被试对于视觉目标的反应。线索有效性和声音呈现的交互作用不显著，$F(1, 33) < 1$，没有证据表明 IOR 效应对于听觉刺激所产生的促进效应产生影响。

3.3 小结

实验 2 的结果显示，被试在线索有效条件下的反应时慢于线索无效条件，IOR 效应出现。声音呈现主效应显著，提前视觉刺激 200 ms 呈现的听觉刺激促进了被试对于视觉目标的反应。但线索有效性和声音呈现之间的交互作用不显著。实验 2 的结果表明，单纯的声音呈现不会与 IOR 效应产生交互，只有听觉刺激与视觉目标发生跨通道对应才会与 IOR 效应产生交互。实验 2 中听觉刺激对于视觉目标的反应促进主要反映了警觉效应，而警觉效应没有和 IOR 效应产生交互。结合以往研究，警觉效应是以一种自上而下的方式增强对于刺激的感知(Kusnir et al., 2011)，而 IOR 以自下而上的方式影响被试的知觉(Berdica et al., 2017; Jia et al., 2019)，由于两种效应作用的通路不同，因此不会产生交互作用。实验 2 支持了实验 1 中的结果的确是 IOR 效应对视听跨通道对应的影响所导致的。

为了进一步探究 IOR 效应对于视听跨通道对应的调节机制，本研究设计了实验 3，实验 3 通过操纵线索和目标间的 SOA 来操纵 IOR 效应的大小，探究 IOR 效应的大小对视听跨通道对应的影响。如果反比效应原则成立，则 SOA 增加所导致的 IOR 效应变小，会使得线索化位置的视听跨通道对应效应减弱，且 IOR 效应对于视听跨通道对应的调节作用减弱。

4　实验 3：不同 SOA 下 IOR 效应对视听跨通道对应的影响

4.1 方法

4.1.1 被试

采用 G*Power 3.1 软件对样本量进行计算。设置 I 类错误的概率 α error prob 为 0.05，检验效能 Power (1 − β err prob)为 0.8，效应量为中等水平(f = 0.25) (Cohen, 1992)，计算被试样本量为 16。招募江苏省高校在校学生 37 名，其中男生 9 名，女生 28 名。剔除被试 3 名，剩余有效被试 34 名，其中男生 9 名，女生 25 名，年龄 19~26 岁。所有被试均为右利手，听力、视力或矫正视力正常，无神经或精神疾病史，无脑部损伤史。被试在实验完成后会获得相应的报酬。

4.1.2 实验仪器与材料

实验 3 在实验 1 的基础上，将线索与目标之间的 SOA 设置为 600 ms 和 1300 ms 两种水平。600 ms 为实验 1 中的线索和目标间的 SOA，1300 ms 是通过延长外源性线索与中央线索之间的间隔时间实现的。其余的实验仪器与实验材料与实验 1 相同。

4.1.3 实验设计和实验流程

实验 3 采用 2 (空间线索有效性：线索有效 vs. 线索无效) × 2 (跨通道对应一致性：对应一致 vs. 对应不一致) × 2 (SOA：600 ms vs. 1300 ms)的被试内设计，因变量为反应时和正确率。正式实验一共 6 组，每组 69 个试次，其中包含 5 个捕获试次，共 414 个试次。被试在进行正式实验前完成 35 个试次的练习实验，整个实验大约进行 50 分钟。其余实验流程均与实验 1 保持一致。

4.2 结果与分析

错误反应、没有反应和反应时异常(100 ms 以内或超出平均反应时正负 3 个标准差)的数据不计入反应时分析，剔除的数据量占总数据量的 1.22%。实验 3 为简单的探测反应任务，被试的平均正确率达到 99%以上，因此不针对正确率进行进一步的统计分析。

4.2.1 反应时

对反应时(见表 1)进行 2 (线索有效性：线索有效 vs. 线索无效) × 2 (跨通道对应一致性：对应一致 vs. 对应不一致) × 2 (SOA：600 ms vs. 1300 ms)的重复测量方差分析。结果显示，线索有效性主效应显著，$F(1, 33) = 89.44$, $p < 0.001$, $\eta_p^2 = 0.73$，线索有效条件下的反应时(355 ms)显著慢于线索无效条件(336 ms), IOR 效应出现。跨通道对应一致性主效应显著，$F(1, 33) = 9.57$, $p = 0.004$, $\eta_p^2 = 0.23$，对应一致条件下的反应时(343 ms)显著快于对应不一致条件(348 ms)，视听跨通道对应效应出现。SOA 主效应不显著，$F(1, 33) < 1$。SOA 和线索有效性的交互作用显著，$F(1, 33) = 6.89$, $p = 0.013$, $\eta_p^2 = 0.17$, SOA 对于 IOR 效应产生了调节作用。简单效应分析表明，当 SOA 为 600 ms 时，线索有效条件下的反应时(356 ms)显著慢于线索无效条件(334 ms), $t(33) = 8.34$, $p < 0.001$, Cohen's $d = 1.43$, 95% CI = [16.33, 26.86]，IOR 效应出现；当 SOA 为 1300 ms 时，线索有效条件下的反应时(354 ms)显著慢于线索无效条件(339 ms)，$t(33) = 8.52$, $p < 0.001$, Cohen's $d = 1.46$, 95% CI = [12.13, 19.74], IOR 效应同样出现。SOA 对于 IOR 效应的调节体现在 SOA 为 600 ms 时 IOR 效应量(22 ms)显著大于 1300 ms 时(16 ms), $t(33) = 2.63$, $p = 0.013$, Cohen's $d = 0.45$, 95% CI = [1.27, 10.05]，随着 SOA 的延长，IOR 效应变小。

表 1 实验 3 中不同条件下的平均反应时($M \pm SD$)

SOA	线索有效条件		线索无效条件	
	对应一致	对应不一致	对应一致	对应不一致
600 ms	350 ± 52	361 ± 51	334 ± 55	335 ± 53
1300 ms	352 ± 53	357 ± 56	336 ± 57	341 ± 57

重要的是，线索有效性、视听跨通道对应一致性和 SOA 三因素交互作用显著，$F(1, 33) = 6.40$, $p = 0.016$, $\eta_p^2 = 0.16$。当 SOA 为 600 ms 时，线索有效性和跨通道对应一致性的交互作用显著，$F(1, 33) = 19.45$, $p < 0.001$, $\eta_p^2 = 0.37$, IOR 效应对于视听跨通道对应产生了调节。简单效应分析表明，当线索有效时，跨通道对应一致条件下的反应时(350 ms)显著快于对应不一致条件(361 ms), $t(33) = 4.97$, $p < 0.001$, Cohen's $d = 0.85$, 95% CI = [−15.36, −6.43]，出现了视听跨通道对应效应；当线索无效时，跨通道对应一致条件下的反应时(334 ms)和对应不一致条件(335 ms)没有显著差异，$t(33) < 1$，没有出现视听跨通道对应效应。当 SOA 为 1300 ms 时，跨通道对应一致性主效应显著，$F(1, 33) = 5.41$, $p = 0.026$, $\eta_p^2 = 0.14$，对应一致条件下的反应时(344 ms)显著快于对应不一致条件(349 ms)，视听跨通道对应效应出现。线索有效性和跨通道对应一致性的交互作用不显著，$F < 1$，无论是线索化位置还是非线索化位置都出现了视听跨通道对应效应，但值得注意的是，此时线索化位置的跨通道对应效应在统计上显著($t(33) = 2.11$, $p = 0.042$, Cohen's $d = 0.36$, 95% CI = [−9.73, −0.19])，而非线索化位置的跨通道对应效应达到边缘显著($t(33) = 1.78$, $p = 0.084$, Cohen's $d = 0.31$, 95% CI = [−9.44, 0.63])，这意味着在长 SOA 下，线索化位置的视听跨通道对应效应更加稳定。

4.2.2 视听跨通道对应效应

首先计算出不同 SOA 条件下不同线索有效性的视听跨通道对应效应量(对应不一致减去对应一致的平均反应时)，再对其进行 2 (SOA：600 ms vs. 1300 ms) × 2 (线索有效性：线索有效 vs. 线索无效)的重复测量方差分析。

结果如图 4 所示。线索有效性主效应显著，$F(1, 33) = 10.45$, $p = 0.003$, $\eta_p^2 = 0.24$，线索有效条件下的跨通道对应效应量(8 ms)显著大于线索无效条件(3 ms)。SOA 主效应不显著，$F(1, 33) < 1$。线索有效性与 SOA 的交互作用显著，$F(1, 33) = 6.40$, $p = 0.016$, $\eta_p^2 = 0.16$。简单效应分析表明，在线索有效条件下，SOA 为 600 ms 时的跨通道对应效应量(11 ms)显著大于 SOA 为 1300 ms 时(5 ms), $t(33) = 2.20$, $p =$

0.035, Cohen's d = 0.38, 95% CI = [0.44, 11.44]; 而在线索无效条件下, SOA 为 600 ms 时的跨通道对应效应量(1 ms)与 1300 ms 时(4 ms)没有显著差异, t(33) = 1.45, p = 0.156。另一个维度的简单效应分析表明, 当 SOA 为 600 ms 时, 线索有效条件下的跨通道对应效应量(11 ms)显著大于线索无效条件(1 ms), t(33) = 4.41, p < 0.001, Cohen's d = 0.76, 95% CI = [5.35, 14.50]; 而当 SOA 为 1300 ms 时, 线索有效条件下的跨通道对应效应量(5 ms)与无效条件下(4 ms)没有显著差异, t(33) < 1。

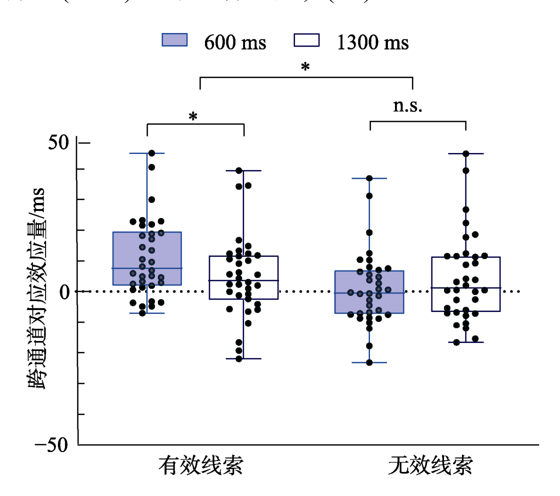


图 4　实验 3 不同 SOA 条件下不同线索有效性的跨通道对应效应量的箱线图

注: 线条自上而下为上限、上四分位数、中位数、下四分位数、下限。散点为个体平均跨通道对应量。*p < 0.05, n.s.代表无显著差异。

4.2.3　IOR 效应

首先计算出不同 SOA 条件下不同线索有效性下的 IOR 效应量, 再对 IOR 效应量进行 2 (SOA: 600 ms vs. 1300 ms) × 2 (跨通道对应一致性: 对应一致 vs. 对应不一致)的重复测量方差分析。

结果显示, SOA 主效应显著, F(1, 33) = 6.89, p = 0.013, η_p^2 = 0.17, SOA 为 600 ms 时的 IOR 效应量(22 ms)显著大于 SOA 为 1300 ms 时(16 ms), 随着 SOA 的延长, IOR 效应量显著降低。跨通道对应一致性主效应显著, F(1, 33) = 10.45, p = 0.003, η_p^2 = 0.24, 跨通道对应一致条件下的 IOR 效应量(16 ms)显著小于对应不一致条件(21 ms)。SOA 与跨通道对应一致性的交互作用显著, F(1, 33) = 6.40, p = 0.016, η_p^2 = 0.16。简单效应分析表明, 当 SOA 为 600 ms 时, 跨通道对应一致条件下的 IOR 效应量(17 ms)显著小于对应不一致条件(27 ms), t(33) = 4.41, p < 0.001, Cohen's d = 0.76, 95% CI = [−14.50, −5.35], 视听跨通道对应产生的促进作用抵消了部分 IOR 效应; 当 SOA 为 1300 ms 时, 跨通道对应一致条件下的 IOR 效应量(16 ms)与对应不一致条件(16 ms)没有显著差异, t(33) < 1。

4.3　小结

实验 3 通过操纵线索和目标之间的 SOA 来操纵 IOR 效应的大小, 探究 IOR 效应的大小对视听跨通道对应的影响。对于 IOR 效应量的分析发现, 随着 SOA 的延长, IOR 效应减弱, 这与前人研究结果相一致(Lupiáñez et al., 1997)。综合统计结果发现, 当 SOA 为 600 ms 时, IOR 效应与视听跨通道对应产生了交互作用, 线索有效条件出现了视听跨通道对应效应, 而线索无效条件下没有出现视听跨通道对应效应, 这与实验 1 的结果相一致。随着 SOA 延长至 1300 ms, 一方面线索有效条件下的视听跨通道对应效应相比于 SOA 为 600 ms 时显著减弱, 另一方面 IOR 效应对于视听跨通道对应的调节作用减弱, 表现为线索有效性与跨通道对应一致性之间交互作用不显著(线索有效条件下的跨通道对应效应与线索无效条件没有显著差异)。当前结果符合反比效应原则(Meredith & Stein, 1983), 即当 SOA 为 1300 ms 时, IOR 效应减弱使得线索化位置视觉目标的感知觉显著性相比于 SOA 为 600 ms 时增强, 更强的视觉信号输入引发了更弱的视听跨通道对应; 同时, 由于 IOR 效应减弱, 线索化位置与非线索化位置视觉目标的感知觉显著性间的差异减少, 使得 IOR 效应对于视听跨通道对应的调节作用减弱, 导致线索化位置和非线索化位置的视听跨通道对应效应没有显著差异。不过, 由于在 1300 ms 条件下 IOR 效应仍然存在, 线索化位置目标感知觉显著性仍然相对较低, 因此统计结果显示线索化位置的视听跨通道对应效应相比于非线索化位置更加稳定。此外, 相比于 600 ms 条件下, SOA 为 1300 ms 时跨通道对应一致条件下的 IOR 效应量与对应不一致条件没有显著差异, 可能是视听跨通道对应效应的减弱导致其对于 IOR 效应的抵抗能力降低。

5　讨论

本研究在空间线索–靶子范式的基础上, 在视觉目标前呈现听觉刺激, 操纵空间线索有效性和视听跨通道对应一致性, 考察 IOR 效应对于视听跨通道对应的影响。实验 1 发现 IOR 效应与视听跨通道对应产生交互, 仅在线索化位置出现了视听跨通道对应效应。实验 1 的结果表明, 视听跨通道对应发生在早期的知觉阶段, 因而会与同样发生在知觉阶

段的 IOR 效应产生交互。实验 2 中，当听觉刺激为单一音调时，IOR 效应没有对听觉刺激对于视觉目标的促进作用产生影响，这意味着只有听觉刺激与视觉刺激发生跨通道对应时才会与 IOR 效应产生交互。此外，实验 2 发现声音刺激所引发的警觉效应不会与 IOR 效应产生交互，进一步支持了实验 1 的结果是由 IOR 效应对于视听跨通道对应的影响所导致的。实验 3 中，研究通过操纵线索与目标之间的 SOA 来操纵 IOR 效应的大小，结果发现随着 IOR 效应的减弱，线索化位置的视听跨通道对应效应变小，且 IOR 效应对于视听跨通道对应的调节作用减弱。这支持了存在于其他多感觉反应增强研究中的反比效应原则也适用于视听跨通道对应。

实验 1 发现 IOR 与视听跨通道对应存在交互作用，根据加因素法反应时的逻辑，两个因素存在交互作用则两个因素作用于同一个加工阶段(Sternberg, 1969)。因此，本研究支持视听跨通道对应发生在与 IOR 效应相同的知觉加工阶段，这与以往一些研究发现一致(Ković et al., 2010; Maeda et al., 2004)。以往部分研究认为视听跨通道对应发生在语义层面上(Gallace & Spence, 2006; Martino & Marks, 1999)，即视、听信息激活了相同的语义编码，导致跨通道对应出现。研究者用“高” “低”两个词的语音代替高、低频率的音调，发现了语音与空间位置出现了和声音频率与空间位置类似的跨通道对应(Gallace & Spence, 2006)。除了直接相关的语义，研究者在音调与亮度的跨通道对应中，将亮度刺激换为“白天” “黑夜”这种语义间接相关的词仍然出现了视听跨通道对应(Martino & Marks, 1999)。目前的研究则表明，视听跨通道对应并不一定要通过语义产生，而是可以发生在纯知觉层面。这符合一些研究发现，一些不用“高” “低”描述音调的种族仍然存在音调与空间位置的跨通道对应(Parkinson et al., 2012)，而一些尚未习得语言的婴儿也表现出视听跨通道对应(Dolscheid et al., 2014; Walker et al., 2010)。当然目前研究结果并不否认语义编码可能在跨通道对应中发挥作用。本研究中的跨通道对应是两个基本刺激特征(音调、空间位置)之间的对应，而声音频率和视觉空间位置本身就存在自然的相关性(Spence, 2011)，在自然界中质量较重的生物的发声频率往往较低(如牛的叫声)，而较重的生物不太可能存在于高位置(如空中)。同时，人类的喉头在发出较低的音调时下降，在发出较高的音调时上升(Parkinson et al., 2012)，这都可能导致音调和空间位置之间存在感知上的对应而不需要通过语义产生。而一些不存在自然相关性的刺激特征之间，或是更复杂的刺激之间的跨通道对应可能需要语义编码来产生(Parise & Spence, 2012)。此外，以往视听跨通道对应的语义研究中，实验任务鼓励被试进行语义编码(Martino & Marks, 1999)，也可能导致了语义在跨通道对应中的作用增加。

以往有研究者使用信号检测论的方法对响度与亮度之间的跨通道对应进行研究，发现视听跨通道对应不影响被试的知觉，只影响被试的决策标准(Marks et al., 2003)，但这一研究与通常所用的视听跨通道对应的范式不同，被试需要判断前后两个刺激的差异，而非探测或辨别单个刺激，因此可能造成了行为反应更多依赖于决策过程。总之，不同刺激之间的视听跨通道对应可能存在不同的机制(Spence, 2011)，不同的实验任务也可能产生了不同的结果。

在实验 1 中，视听跨通道对应仅在线索化位置出现，这与部分视听觉整合的研究结果相类似。以往视听觉整合的研究中，有研究发现 IOR 效应发生时，线索化位置的视听觉整合效应更大(唐晓雨 等, 2020)，研究者使用反比效应原则对于结果进行解释，即视、听感觉信号的输入强度会影响视听觉整合效应的大小。根据反比效应原则(Otto et al., 2013)，视觉或听觉感觉信号输入强度弱时，视听觉整合效应更大；而视觉或听觉感觉信号输入强度强时，视听觉整合效应更小(Senkowski et al., 2011)。反比效应原则最初是在动物的单神经元研究中发现的(Meredith & Stein, 1983; Stein & Meredith, 1993)，行为和神经上的研究表明人类被试的多感觉反应增强仍然遵循这一原则(Rach et al., 2011; Senkowski et al., 2011; van de Rijt et al., 2019)。反比效应原则本质上反映的是更弱的冗余性信息引发了更强的整合。本研究的视听跨通道对应中，尽管视、听刺激在形式上是非冗余性的，但有研究者认为这些刺激处于心理坐标中的相同一端，如高音和高空间位置都意味着相比于低音和低空间位置更“高”的一端(McCormick et al., 2018)，因此可能存在共同的神经编码，产生了冗余性，因而反比效应原则可能也适用于视听跨通道对应。当 IOR 效应发生时，线索化位置的目标感知觉显著性降低(Slagter et al., 2016)，视觉信号输入减弱，因而产生了稳定的视听跨通道对应效应，而非线索化位置相对高的目标感知觉显著性没有产生视听跨通道对应效应。

实验 3 的结果发现, 随着 IOR 效应减弱, 线索化位置的视听跨通道对应效应也相应减弱, 同时, IOR 效应对于视听跨通道对应的调节作用减弱, 表现为线索有效性与跨通道对应一致性之间的交互作用不显著(线索有效条件下的跨通道对应效应与线索无效条件没有显著差异), 直接验证了反比效应原则对于视听跨通道对应的适用性。在多感觉反应增强的研究中, 反比效应原则不仅存在于无意义的视、听刺激整合中(Senkowski et al., 2011), 也存在于多感觉输入下的言语知觉中(van de Rijt et al., 2019), 研究者发现无论是听觉输入还是视觉输入, 更难以感知的单词输入引发了更大的多感觉反应增强。目前的研究结果支持视听跨通道对应的发生也符合反比效应原则, 扩展了反比效应原则在多感觉反应增强领域的适用范围。

研究同样发现, 视听跨通道对应也对 IOR 效应产生了影响。具体来讲, 当 SOA 为 600 ms 时, 无论在实验 1 还是实验 3 中, 跨通道对应一致条件下的 IOR 效应量都显著小于对应不一致条件, 这是由于听觉刺激与视觉目标之间的跨通道对应增加了视觉目标的感知觉显著性, 部分抵消了 IOR 效应导致的目标感知觉显著性降低。而随着 SOA 延长, 跨通道对应一致条件下的 IOR 效应和不一致条件没有显著差异。这是由于长 SOA 下, 视听跨通道对应效应减弱, 进而抵抗 IOR 效应的能力也随之降低, 从而使得不同对应一致性条件下的 IOR 效应无显著差异。

目前研究是第一个探究 IOR 效应与视听跨通道对应之间关系的研究。因此本研究需要确保实验中的交互作用的确是由 IOR 效应对于视听跨通道对应进行调节所产生的。实验 2 在实验 1 其他条件保持不变的基础上, 通过操纵听觉刺激是否出现, 来探究单纯的声音呈现在本研究中的影响。由于视听跨通道对应是一种相对的映射(Chiou & Rich, 2012), 需要两个有相对高、低关系的音调才能与高、低位置相对应, 因此单独的纯音不会引发视听跨通道对应。结果发现 IOR 效应与听觉刺激是否出现不存在交互作用。实验 2 的结果一方面验证了单纯的声音呈现不会与 IOR 效应产生交互, 只有听觉刺激与视觉刺激发生跨通道对应才会与 IOR 效应产生交互。另一方面表明听觉刺激所产生的警觉效应没有和 IOR 效应产生交互, 排除了警觉效应对于实验结果可能存在的混淆。有研究表明, 听觉刺激产生的警觉效应会提高被试的感知能力(Kusnir et al., 2011), 并与目标的感知觉显著性产生交互作用(Botta et al., 2017), 表现为感知显著性更高的目标产生了更大的警觉效应。来自行为和神经上的证据表明, 警觉效应也会与空间注意产生交互(Botta et al., 2014; 2017), 警觉效应相关的唤醒系统(arousal system)会与注意系统存在补偿机制(Fischer et al., 2008; Portas et al., 1998)。但实验 2 结果表明, 在目前研究中警觉效应与外源性线索引发的 IOR 效应缺乏交互作用。这一方面可能由于, 以往研究所比较的是阈上、阈下和阈刺激与警觉效应之间的相互作用(Botta et al., 2017; Chica et al., 2016), 而本研究中视觉目标出现在线索化位置和非线索化位置都完全可见, IOR 效应发生所导致的目标感知显著性的差异不足以引发警觉效应的调节。另一方面可能是警觉效应和 IOR 效应沿着不同的神经通路独立发生。尽管有研究表明, 警觉效应和 IOR 效应都影响刺激的感知觉显著性(Botta et al., 2014; Prime & Jolicoeur, 2009), 且都与额顶网络的激活有关(Bourgeois et al., 2012; Kusnir et al., 2011), 但警觉效应是以一种自上而下的方式增加对于刺激的感知(Kusnir et al., 2011), 即通过激活个体的额顶网络以放大输入刺激的强度; 而外源性线索引发的 IOR 以自下而上的方式影响被试的知觉(Berdica et al., 2017; Jia et al., 2019), 通过调节输入刺激的强度, 影响早期视觉区域对于额顶网络的投射(Botta et al., 2014)。当然, 警觉效应与 IOR 效应之间的关系需要神经水平上的进一步的探究。目前的研究排除了警觉效应对于结果的影响, 支持了音调与空间位置之间的视听跨通道对应发生在知觉水平的观点。

6　结论

IOR 效应对于视听跨通道对应产生了调节作用, IOR 发生时线索化位置出现了稳定的视听跨通道对应效应, 而非线索化位置没有出现。听觉刺激引发的警觉效应没有与 IOR 效应产生交互。随着 IOR 效应的减弱, 线索化位置的视听跨通道对应效应也随之减弱, 且 IOR 效应对于视听跨通道对应的调节作用减弱。当前研究结果支持了听觉音调与视觉空间位置之间的视听跨通道对应发生在知觉水平, 且视听跨通道对应的发生符合反比效应原则。

参　考　文　献

Berdica, E., Gerdes, A. B. M., & Alpers, G. W. (2017). A comprehensive look at phobic fear in inhibition of return: Phobia-related spiders as cues and targets. *Journal of*

Behavior Therapy and Experimental Psychiatry, 54, 158−164.

Botta, F., Lupiáñez, J., & Chica, A. B. (2014). When endogenous spatial attention improves conscious perception: Effects of alerting and bottom-up activation. *Consciousness and Cognition, 23*, 63−73.

Botta, F., Ródenas, E., & Chica, A. B. (2017). Target bottom-up strength determines the extent of attentional modulations on conscious perception. *Experimental Brain Research, 235*(7), 2109−2124.

Bourgeois, A., Chica, A. B., Migliaccio, R., Thiebaut de Schotten, M., & Bartolomeo, P. (2012). Cortical control of inhibition of return: Evidence from patients with inferior parietal damage and visual neglect. *Neuropsychologia, 50*(5), 800−809.

Brunel, L., Carvalho, P. F., & Goldstone, R. L. (2015). It does belong together: Cross-modal correspondences influence cross-modal integration during perceptual learning. *Frontiers in Psychology, 6*, 358.

Brunetti, R., Indraccolo, A., Del Gatto, C., Spence, C., & Santangelo, V. (2018). Are crossmodal correspondences relative or absolute? Sequential effects on speeded classification. *Attention, Perception & Psychophysics*, *80*(2), 527−534.

Chica, A. B., Bayle, D. J., Botta, F., Bartolomeo, P., & Paz-Alonso, P. M. (2016). Interactions between phasic alerting and consciousness in the fronto-striatal network. *Scientific Reports, 6*, 31868.

Chica, A. B., Lasaponara, S., Chanes, L., Valero-Cabré, A., Doricchi, F., Lupiáñez, J., & Bartolomeo, P. (2011). Spatial attention and conscious perception: The role of endogenous and exogenous orienting. *Attention, Perception & Psychophysics, 73*(4), 1065−1081.

Chiou, R., & Rich, A. N. (2012). Cross-modality correspondence between pitch and spatial location modulates attentional orienting. *Perception, 41*(3), 339−353.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155−159.

Dolscheid, S., Hunnius, S., Casasanto, D., & Majid, A. (2014). Prelinguistic infants are sensitive to space-pitch associations found across cultures. *Psychological Science, 25*(6), 1256−1261.

Erdfelder, E., Auer, T. S., Hilbig, B. E., Aßfalg, A., Moshagen, M., & Nadarevic, L. (2009). Multinomial processing tree models: A review of the literature. *Natural Science Journal of Harbin Normal University, 217*(3), 108−124.

Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision, 10*(1), 1−12.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175−191.

Fischer, T., Langner, R., Birbaumer, N., & Brocke, B. (2008). Arousal and attention: Self-chosen stimulation optimizes cortical excitability and minimizes compensatory effort. *Journal of Cognitive Neuroscience, 20*(8), 1443−1453.

Frassinetti, F., Bolognini, N., & Làdavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental Brain Research, 147*(3), 332−343.

Gallace, A., & Spence, C. (2006). Multisensory synesthetic interactions in the speeded classification of visual size. *Perception & Psychophysics, 68*(7), 1191−1203.

Getz, L. M., & Kubovy, M. (2018). Questioning the automaticity of audiovisual correspondences. *Cognition, 175*, 101−108.

Hopfinger, J. B., & Mangun, G. R. (2001). Tracking the influence of reflexive attention on sensory and cognitive processing. *Cognitive, Affective & Behavioral Neuroscience, 1*(1), 56−65.

Jia, L., Wang, J., Zhang, K., Ma, H., & Sun, H. J. (2019). Do emotional faces affect inhibition of return? An ERP study. *Frontiers in Psychology, 10*(721), 1−8.

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Science, 4*(4), 138−147.

Ković, V., Plunkett, K., & Westermann, G. (2010). The shape of words in the brain. *Cognition, 114*(1), 19−28.

Kusnir, F., Chica, A. B., Mitsumasu, M. A., & Bartolomeo, P. (2011). Phasic auditory alerting improves visual conscious perception. *Consciousness and Cognition, 20*(4), 1201−1210.

Lupiáñez, J., Milán, E. G., Tornay, F. J., Madrid, E., & Tudela, P. (1997). Does IOR occur in discrimination tasks? Yes, it does, but later. *Perception & Psychophysics, 59*(8), 1241−1254.

Maeda, F., Kanai, R., & Shimojo, S. (2004). Changing pitch induced visual motion illusion. *Current Biology, 14*(23), 990−991.

Maimon, N. B., Lamy, D., & Eitan, Z. (2020). Crossmodal correspondence between tonal hierarchy and visual brightness: Associating syntactic structure and perceptual dimensions across modalities. *Multisensory Research, 33*(8), 805−836.

Marks, L. E., Ben-Artzi, E., & Lakatos, S. (2003). Cross-modal interactions in auditory and visual discrimination. *International Journal of Psychophysiology, 50*(1−2), 125−145.

Martino, G., & Marks, L. E. (1999). Perceptual and linguistic interactions in speeded classification: Tests of the semantic coding hypothesis. *Perception, 28*(7), 903−923.

McCormick, K., Lacey, S., Stilla, R., Nygaard, L. C., & Sathian, K. (2018). Neural basis of the crossmodal correspondence between auditory pitch and visuospatial elevation. *Neuropsychologia, 112*, 19−30.

McCracken, H. S., Murphy, B. A., Glazebrook, C. M., Burkitt, J. J., Karellas, A. M., & Yielder, P. C. (2019). Audiovisual multisensory integration and evoked potentials in young adults with and without attention-deficit/hyperactivity disorder. *Frontiers in Human Neuroscience, 13*, 95.

Meredith, M. A., & Stein, B. E. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science, 221*(4608), 389−391.

Noesselt, T., Rieger, J. W., Schoenfeld, M. A., Kanowski, M., Hinrichs, H., Heinze, H. J., & Driver, J. (2007). Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. *Journal of Neuroscience, 27*(42), 11431−11441.

Otto, T. U., Dassy, B., & Mamassian, P. (2013). Principles of multisensory behavior. *Journal of Neuroscience, 33*(17), 7463−7474.

Parise, C., & Spence, C. (2008). Synesthetic congruency modulates the temporal ventriloquism effect. *Neuroscience Letters, 442*(3), 257−261.

Parise, C., & Spence, C. (2012). Audiovisual crossmodal correspondences and sound symbolism: A study using the implicit association test. *Experimental Brain Research, 220*(3−4), 319−333.

Parkinson, C., Kohler, P. J., Sievers, B., & Wheatley, T. (2012). Associations between auditory pitch and visual elevation do not depend on language: Evidence from a remote population. *Perception, 41*(7), 854−861.

Peng, X., Chang, R., Li, Q., Wang, A., & Tang, X. (2019). Visually induced inhibition of return affects the audiovisual integration under different SOA conditions. *Acta Psychologica*

Sinica, 51(7), 759−771.
[彭姓, 常若松, 李奇, 王爱君, 唐晓雨. (2019). 不同 SOA 下视觉返回抑制对视听觉整合的调节作用. *心理学报, 51*(7), 759−771.]
Portas, C. M., Rees, G., Howseman, A. M., Josephs, O., Turner, R., & Frith, C. D. (1998). A specific role for the thalamus in mediating the interaction of attention and arousal in humans. *The Journal of Neuroscience, 18*(21), 8979−8989.
Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. G. Bowhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 531−556). Erlbaum.
Prime, D. J., & Jolicoeur, P. (2009). On the relationship between occipital cortex activity and inhibition of return. *Psychophysiology, 46*(6), 1278−1287.
Rach, S., Diederich, A., & Colonius, H. (2011). On quantifying multisensory interaction effects in reaction time and detection rate. *Psychological Research, 75*(2), 77−94.
Redden, R. S., Maclnnes, W. J., & Klein, R. M. (2021). Inhibition of return: An information processing theory of its natures and significance. *Cortex, 135*, 30−48.
Santangelo, V., Ho, C., & Spence, C. (2008). Capturing spatial attention with multisensory cues. *Psychonomic Bulletin & Review, 15*(2), 398−403.
Satel, J., Hilchey, M. D., Wang, Z. G., Story, R., & Klein, R. M. (2013). The effects of ignored versus foveated cues upon inhibition of return: An event-related potential study. *Attention, Perception, & Psychophysics, 75*(1), 29−40.
Senkowski, D., Saint-Amour, D., Höfle, M., & Foxe, J. J. (2011). Multisensory interactions in early evoked brain activity follow the principle of inverse effectiveness. *Neuroimage, 56*(4), 2200−2208.
Slagter, H. A., Prinssen, S., Reteig, L. C., & Mazaheri, A. (2016). Facilitation and inhibition in attention: Functional dissociation of pre-stimulus alpha activity, P1, and N1 components. *Neuroimage, 125*(6), 25−35.
Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics, 73*(4), 971−995.
Spence, C. (2013). Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule. *Annals of the New York Academy of Sciences, 1296*(1), 31−49.
Spence, C. (2019). On the Relative nature of (pitch-based) crossmodal correspondences. *Multisensory Research, 32*(3), 235−265.
Starke, J., Ball, F., Heinze, H. J., & Noesselt, T. (2017). The spatio-temporal profile of multisensory integration. *The European Journal of Neuroscience, 51*(5), 1210−1223.
Stein, B. E., & Meredith, M. A. (1993). The merging of the senses. *Journal of Cognitive Neuroscience, 5*(3), 373−374.
Stein, B. E., Meredith, M. A., Huneycutt, W. S., & McDade, L. (1989). Behavioral indices of multisensory integration: Orientation to visual cues is affected by auditory stimuli. *Journal of Cognitive Neuroscience, 1*(1), 12−24.
Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: Current issues from the perspective of the single neuron. *Nature Reviews Neuroscience, 9*(4), 255−266.
Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276−315.
Talsma, D., & Woldorff, M. G. (2005). Selective attention and multisensory integration: Multiple phases of effects on the evoked brain activity. *Journal of Cognitive Neuroscience, 17*(7), 1098−1114.
Tang, X., Gao, Y., Yang, W., Ren, Y., Wu, J., Zhang, M., & Wu, Q. (2019). Bimodal-divided attention attenuates visually induced inhibition of return with audiovisual targets. *Experimental Brain Research, 237*(4), 1093−1107.
Tang, X., Sun, J., & Peng, X. (2020). The effect of bimodal divided attention on inhibition of return with audiovisual targets. *Acta Psychologica Sinica, 52*(3), 257−268.
[唐晓雨, 孙佳影, 彭姓. (2020). 双通道分配性注意对视听觉返回抑制的影响. *心理学报, 52*(3), 257−268.]
Tang, X., Wu, J., & Shen, Y. (2016). The interactions of multisensory integration with endogenous and exogenous attention. *Neuroscience and Biobehavioral Reviews, 61*, 208−224.
van de Rijt, L. P. H., Roye, A., Mylanus, E. A. M., van Opstal, A. J., & van Wanrooij, M. M. (2019). The principle of inverse effectiveness in audiovisual speech perception. *Frontiers in Human Neuroscience, 13*, 335.
van der Stoep, N., Spence, C., Nijboer, T. C., & van der Stigchel, S. (2015). On the relative contributions of multisensory integration and crossmodal exogenous spatial attention to multisensory response enhancement. *Acta Psychologica, 162*, 20−28.
van der Stoep, N., van der Stigchel, S., & Nijboer, T. C. W. (2015). Exogenous spatial attention decreases audiovisual integration. *Attention Perception & Psychophysics, 77*(1), 464−482.
van der Stoep, N., van der Stigchel, S., Nijboer, T. C. W., & Spence, C. (2016). Visually induced inhibition of return affects the integration of auditory and visual information. *Perception, 46*(1), 6−17.
Walker, P. (2012). Cross-sensory correspondences and cross talk between dimensions of connotative meaning: Visual angularity is hard, high-pitched, and bright. *Attention, Perception & Psychophysics, 74*(8), 1792−1809.
Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological Science, 21*(1), 21−25.
Wiegand, I., & Sander, M. C. (2019). Cue-related processing accounts for age differences in phasic alerting. *Neurobiology of Aging, 79*, 93−100.
Zeljko, M., Kritikos, A., & Grove, P. M. (2019). Lightness/pitch and elevation/pitch crossmodal correspondences are low-level sensory effects. *Attention, Perception & Psychophysics, 81*(5), 1609−1623.

Effect of inhibition of return on audiovisual cross-modal correspondence

ZU Guangyao[1], LI Shuqi[1], ZHANG Tianyang[2], WANG Aijun[1], ZHANG Ming[1]

([1] *Department of Psychology, Research Center for Psychology and Behavioral Sciences, Soochow University, Suzhou* 215123, *China*)

([2] *School of Public Health, Soochow University, Suzhou* 215123, *China*)

Abstract

Different dimensions of visual and auditory stimuli can map to each other to influence human behavioral responses, a phenomenon known as audiovisual cross-modal correspondence. A common audiovisual cross-modal correspondence is between auditory tones and visual spatial locations, with individuals tending to map high-pitched sounds to high spatial location and low-pitched sounds to low spatial location. When a high-pitch sound is accompanied or preceded by a visual stimulus, the participants respond faster to visual stimuli presented in the high spatial location than to visual stimuli presented in the low spatial location, and vice versa. Researchers have different views on the level at which audiovisual cross-modal correspondence occurs. Some argue that audiovisual cross-modal correspondence occurs at the perceptual level, increasing the perceptual saliency of the stimulus, while others argue that audiovisual cross-modal correspondence occurs at a later semantic or decision level. As inhibition of return (IOR) in the attentional system can affect human perception, this study used a cue-target paradigm to explore the interaction between IOR and audiovisual cross-modal correspondence to elucidate the occurrence level and mechanism of audiovisual cross-modal correspondence. Audiovisual cross-modal correspondence between auditory tones and visual spatial locations was expected to occur at the perceptual level and therefore would be subject to the IOR effect occurring at the same processing level.

The present study consisted of 3 experiments. Experiment 1 had a 2 × 2 within-subjects design; we manipulated the spatial cue validity (valid cue vs. invalid cue) and audiovisual cross-modal correspondence (congruent vs. incongruent). During the experiment, a fixation point was first presented in the middle of the screen for 750 ms. The box above or below the fixation point was then bolded for 50 ms, but this cue was not predictive of the spatial location of the target. After a time interval of 250 ms, a fixation point was presented in bold as a central cue. A central cue is commonly used in spatial IOR research, as it facilitates stable occurrence of IOR. The central cue was presented for 50 ms, and then the auditory stimulus (either high or low pitch) was presented for 50 ms. After a 200-ms interval, the visual target was presented for 100 ms in the box above or below the fixation point. The participants were instructed to perform a detection task for the presence of a visual target. The experimental design and procedure of Experiment 2 were identical to those of Experiment 1, except that the sound presented before the visual target was a single tone that was present or absent. Experiment 3 had a 2 × 2 × 2 within-subjects design. Experiment 3 added a factor to Experiment 1, namely, stimulus onset asynchrony (SOA) between the cue and the target (600 ms vs. 1300 ms).

In all three experiments, the overall accuracy (ACC) was very high; thus, no further statistical analysis was conducted for the ACC. In terms of reaction time (RT), the results of Experiment 1 showed that both spatial IOR and audiovisual cross-modal correspondence occurred. Importantly, there was an interaction between spatial cue validity and audiovisual cross-modal correspondence. Specifically, when the cue was valid, audiovisual cross-modal correspondence occurred, and when the cue was invalid, there was no audiovisual cross-modal correspondence. The results of Experiment 2 showed that the interaction between cue validity and sound presentation was not significant, and there was no evidence that IOR influenced the sound-induced facilitation effect. The results of Experiment 3 showed that the interaction among spatial cue validity, cross-modal correspondence congruency, and SOA was significant. Specifically, at an SOA of 600 ms, the interaction between spatial cue validity and cross-modal correspondence congruency was significant. When the cue was

valid, audiovisual cross-modal correspondence occurred, and when the cue was invalid, there was no audiovisual cross-modal correspondence. At an SOA of 1300 ms, the interaction between cue validity and cross-modal correspondence congruency was not significant, and cross-modal correspondence occurred in both valid-cue and invalid-cue conditions. The results of the analysis of the IOR effect showed that the IOR effect under the 600-ms SOA condition (22 ms) was significantly larger than that under the 1300-ms SOA condition (16 ms). As the SOA increased, the IOR effect size decreased.

In conclusion, the present results suggested that the IOR effect, occurring at the perceptual level, moderated audiovisual cross-modal correspondence. When the IOR effect occurred, audiovisual cross-modal correspondence occurred in the cued location, but not in the non-cued location. The alerting effect induced by the sound did not interact with IOR. With the weakening of the IOR effect, the audiovisual cross-modal correspondence in the cued location decreased, and the moderating effect of the IOR effect on audiovisual cross-modal correspondence weakened. The present results support that the audiovisual cross-modal correspondence between auditory tones and visual spatial locations occurs at the perceptual level, and the occurrence of audiovisual cross-modal correspondence conforms to the principle of inverse effectiveness.

Keywords　audiovisual cross-modal correspondence, inhibition of return, cue-target paradigm, alerting effect